

Bloomberg describes BloombergGPT, a 50-billion-parameter model trained on financial data

Source copy kept by wheresthe.ai for the story "Bloomberg describes BloombergGPT, a 50-billion-parameter model trained on financial data" (<https://wheresthe.ai/d/bloomberg-bloomberggpt-paper-2023-03/>).

Source: <https://arxiv.org/abs/2303.17564v1>

Text saved 18 Oct 2026 (the page as published).

SHA-256 of the saved text: 810a0f7aa217b0b9e6a4e31116a8b40d45c06e18f35813202371fe4f4f940f07

-->

Computer Science > Machine Learning

arXiv:2303.17564v1 (cs)

[Submitted on 30 Mar 2023 (this version), latest version 21 Dec 2023 (v3)]

Title: BloombergGPT: A Large Language Model for Finance

Authors: Shijie Wu , Ozan Irsoy , Steven Lu , Vadim Dabravolski , Mark Dredze , Sebastian Gehrmann , Prabhanjan Kambadur , David Rosenberg , Gideon Mann View a PDF of the paper titled BloombergGPT: A Large Language Model for Finance, by Shijie Wu and 8 other authors

View PDF HTML (experimental)

Abstract: The use of NLP in the realm of financial technology is broad and complex, with applications ranging from sentiment analysis and named entity recognition to question answering. Large Language Models (LLMs) have been shown to be effective on a variety of tasks; however, no LLM specialized for the financial domain has been reported in literature. In this work, we present BloombergGPT, a 50 billion parameter language model that is trained on a wide range of financial data. We construct a 363 billion token dataset based on Bloomberg's extensive data sources, perhaps the largest domain-specific dataset yet, augmented with 345 billion tokens from general purpose datasets. We validate BloombergGPT on standard LLM benchmarks, open financial benchmarks, and a suite of internal benchmarks that most accurately reflect our intended usage. Our mixed dataset training leads to a model that outperforms existing models on financial tasks by significant margins without sacrificing performance on general LLM benchmarks. Additionally, we explain our modeling choices, training process, and evaluation methodology. As a next step, we plan to release training logs (Chronicles) detailing our experience in training BloombergGPT.

Subjects:

Machine Learning (cs.LG) ; Artificial Intelligence (cs.AI); Computation and Language (cs.CL); General Finance (q-fin.GN)

Cite as: arXiv:2303.17564 [cs.LG]

(or arXiv:2303.17564v1 [cs.LG] for this version)

<https://doi.org/10.48550/arXiv.2303.17564>

Focus to learn more

arXiv-issued DOI via DataCite

Submission history From: Sebastian Gehrmann [view email] [v1] Thu, 30 Mar 2023 17:30:36 UTC (7,463 KB)

[v2] Tue, 9 May 2023 16:06:35 UTC (1,328 KB)

[v3] Thu, 21 Dec 2023 06:21:11 UTC (1,328 KB)

Full-text links: Access Paper:

View a PDF of the paper titled BloombergGPT: A Large Language Model for Finance, by Shijie Wu and 8 other authors View PDF HTML (experimental) TeX Source

[view license](#)

Current browse context:

[cs.LG](#)

[< prev](#)

[| next >](#)

[new](#) | [recent](#) | [2023-03](#)

Change to browse by:

[cs](#) [cs.AI](#) [cs.CL](#) [q-fin](#) [q-fin.GN](#)

References & Citations

[NASA ADS](#) [Google Scholar](#)

[Semantic Scholar](#)

[export BibTeX citation](#) [Loading...](#)

[BibTeX formatted citation](#)

[×](#)

[loading...](#)

Data provided by:

[Bookmark](#)

[Bibliographic Tools](#)

[Bibliographic and Citation Tools](#)

[Bibliographic Explorer Toggle](#)

[Bibliographic Explorer \(What is the Explorer? \)](#)

[Connected Papers Toggle](#)

[Connected Papers \(What is Connected Papers? \)](#)

[Litmaps Toggle](#)

[Litmaps \(What is Litmaps? \)](#)

[scite.ai Toggle](#)

[scite Smart Citations \(What are Smart Citations? \)](#)

[Code, Data, Media](#)

[Code, Data and Media Associated with this Article](#)

[alphaXiv Toggle](#)

[alphaXiv \(What is alphaXiv? \)](#)

[Links to Code Toggle](#)

[CatalyzeX Code Finder for Papers \(What is CatalyzeX? \)](#)

[DagsHub Toggle](#)

[DagsHub \(What is DagsHub? \)](#)

[GotitPub Toggle](#)

[Gotit.pub \(What is GotitPub? \)](#)

[Huggingface Toggle](#)

[Hugging Face \(What is Huggingface? \)](#)

[ScienceCast Toggle](#)

[ScienceCast \(What is ScienceCast? \)](#)

[Demos](#)

[Demos](#)

[Replicate Toggle](#)

[Replicate \(What is Replicate? \)](#)

[Spaces Toggle](#)

[Hugging Face Spaces \(What is Spaces? \)](#)

[Spaces Toggle](#)

[TXYZ.AI \(What is TXYZ.AI? \)](#)

[Related Papers](#)

[Recommenders and Search Tools](#)

[Link to Influence Flower](#)

[Influence Flower \(What are Influence Flowers? \)](#)

[Core recommender toggle](#)

[CORE Recommender \(What is CORE? \)](#)

[IArxiv recommender toggle](#)

[IArxiv Recommender \(What is IArxiv? \)](#)

[Author](#)

[Venue](#)

[Institution](#)

[Topic](#)

[About arXivLabs](#)

arXivLabs: experimental projects with community collaborators

arXivLabs is a framework that allows collaborators to develop and share new arXiv features directly on our website.

Both individuals and organizations that work with arXivLabs have embraced and accepted our values of openness, community, excellence, and user data privacy. arXiv is committed to these values and only works with partners that adhere to them.

Have an idea for a project that will add value for arXiv's community? [Learn more about arXivLabs](#) .

[Which authors of this paper are endorsers? | Disable MathJax \(What is MathJax? \)](#)